

Supplementary Material

Shotgun Mitogenomics Provides a Reference Phylogenetic Framework and Timescale for Living Xenarthrans

Gillian C. Gibb^{1,2}, Fabien L. Condamine^{1,3,4}, Melanie Kuch⁵, Jacob Enk⁵, Nadia Moraes-Barros^{6,7}, Mariella Superina⁸, Hendrik N. Poinar^{5,*} and Frédéric Delsuc^{1,*}

Supplementary Table S1: Detailed results of the PartitionFinder analysis.

Supplementary Table S2: Detailed results of diversification analyses.

Supplementary Figure S1: Complete Bayesian consensus tree.

Supplementary Figure S2: Complete Bayesian chronogram.

Supplementary Figure S3: Detailed BAMM results.

Supplementary Table S1: Detailed results of the PartitionFinder analysis.

Subset	Best Model	Subset Partitions
1	GTR+I+G	12S rRNA, 16S rRNA, tRNAs, ATP6_p1, ATP8_p1, ATP8_p2, CO1_p1, CO2_p1, CO3_p1, CYTB_p1, ND1_p1, ND2_p1, ND3_p1, ND4L_p1, ND4_p1, ND5_p1
2	GTR+I+G	ATP6_p2, CO1_p2, CO2_p2, CO3_p2, CYTB_2, ND1_p2, ND2_p2, ND3_p2, ND4L_p2, ND4_p2, ND5_p2
3	GTR+I+G	ATP6_p3, ATP8_p3, CO1_p3, CO2_p3, CO3_p3, CYTB_p3, ND1_p3, ND2_p3, ND3_p3, ND4L_p3, ND4_p3, ND5_p3
4	GTR+I+G	ND6_p1, ND6_p2, ND6_p3

Settings used

```
alignment      : ./40Xen.phy
branchlengths  : unlinked
models        : GTR+I, GTR+G, TrNef+G, TrN+G, TrN+I, TrNef+I, HKY, K80, HKY+I+G,
K80+I, SYM+G, SYM+I, K80+G, TrNef, GTR, TrNef+I+G, HKY+G, SYM+I+G, TrN, HKY+I, SYM,
GTR+I+G, K80+I+G, TrN+I+G
model_selection : bic
search         : greedy
```

Best partitioning scheme

Scheme Name: step_38
Scheme lnL: -161908.30861
Scheme BIC: 327160.98658
Number of params: 348
Number of sites: 14917
Number of subsets: 4

Subset | Best Model | Subset Partitions | Subset Sites

1 GTR+I+G 12S, 16S, atp6_1, atp8_1, atp8_2, co1_1, co2_1, co3_1, cyt b_1, nd1_1, nd2_1, nd3_1, nd4L_1, nd4_1, nd5_1, tRNAs 1-880, 881-2244, 2245-2922\3, 2923-3099\3, 2924-3099\3, 3100-4638\3, 4639-5319\3, 5320-6102\3, 6103-7236\3, 7237-8187\3, 8188- 9219\3, 9220-9564\3, 9565-10941\3, 10942-11235\3, 11236-13023\3, 13537-14917
2 GTR+I+G atp6_2, co1_2, co2_2, co3_2, cyt b_2, nd1_2, nd2_2, nd3_2, nd4L_2, nd4_2, nd5_2 2246-2922\3, 3101-4638\3, 4640-5319\3, 5321-6102\3, 6104-7236\3, 7238- 8187\3, 8189-9219\3, 9221-9564\3, 9566-10941\3, 10943-11235\3, 11237-13023\3
3 GTR+I+G atp6_3, atp8_3, co1_3, co2_3, co3_3, cyt b_3, nd1_3, nd2_3, nd3_3, nd4L_3, nd4_3, nd5_3 2247-2922\3, 2925-3099\3, 3102-4638\3, 4641-5319\3, 5322- 6102\3, 6105-7236\3, 7239-8187\3, 8190-9219\3, 9222-9564\3, 9567-10941\3, 10944- 11235\3, 11238-13023\3
4 GTR+I+G nd6_1, nd6_2, nd6_3 13024-13536\3, 13025-13536\3, 13026-13536\3

Scheme Description in PartitionFinder format

Scheme_step_38 = (12S, 16S, atp6_1, atp8_1, atp8_2, co1_1, co2_1, co3_1, cyt b_1, nd1_1,
nd2_1, nd3_1, nd4L_1, nd4_1, nd5_1, tRNAs) (atp6_2, co1_2, co2_2, co3_2, cyt b_2, nd1_2,

nd2_2, nd3_2, nd4L_2, nd4_2, nd5_2) (atp6_3, atp8_3, co1_3, co2_3, co3_3, cytB_3, nd1_3, nd2_3, nd3_3, nd4L_3, nd4_3, nd5_3) (nd6_1, nd6_2, nd6_3);

RaxML-style partition definitions

DNA, p1 = 1-880, 881-2244, 2245-2922\3, 2923-3099\3, 2924-3099\3, 3100-4638\3, 4639-5319\3, 5320-6102\3, 6103-7236\3, 7237-8187\3, 8188-9219\3, 9220-9564\3, 9565-10941\3, 10942-11235\3, 11236-13023\3, 13537-14917
DNA, p2 = 2246-2922\3, 3101-4638\3, 4640-5319\3, 5321-6102\3, 6104-7236\3, 7238-8187\3, 8189-9219\3, 9221-9564\3, 9566-10941\3, 10943-11235\3, 11237-13023\3
DNA, p3 = 2247-2922\3, 2925-3099\3, 3102-4638\3, 4641-5319\3, 5322-6102\3, 6105-7236\3, 7239-8187\3, 8190-9219\3, 9222-9564\3, 9567-10941\3, 10944-11235\3, 11238-13023\3
DNA, p4 = 13024-13536\3, 13025-13536\3, 13026-13536\3

Supplementary Table S2: Detailed results of diversification analyses.

(a) Paleoenvironmental-dependence approach (RPANDA; Condamine et al. 2013). Values are means and standard errors calculated over 100 trees taken randomly from the Bayesian dating analyses. NP, number of parameters; logL, log-likelihood; λ , speciation rate; α , variation of the speciation rate according the temperature; μ , extinction rate; and β , variation of the extinction rate according the temperature. The best-fitting model is the model with speciation varying (exponentially) negatively with temperature. The linear variation provides a slightly lower fit, and depicts the same response to temperature.

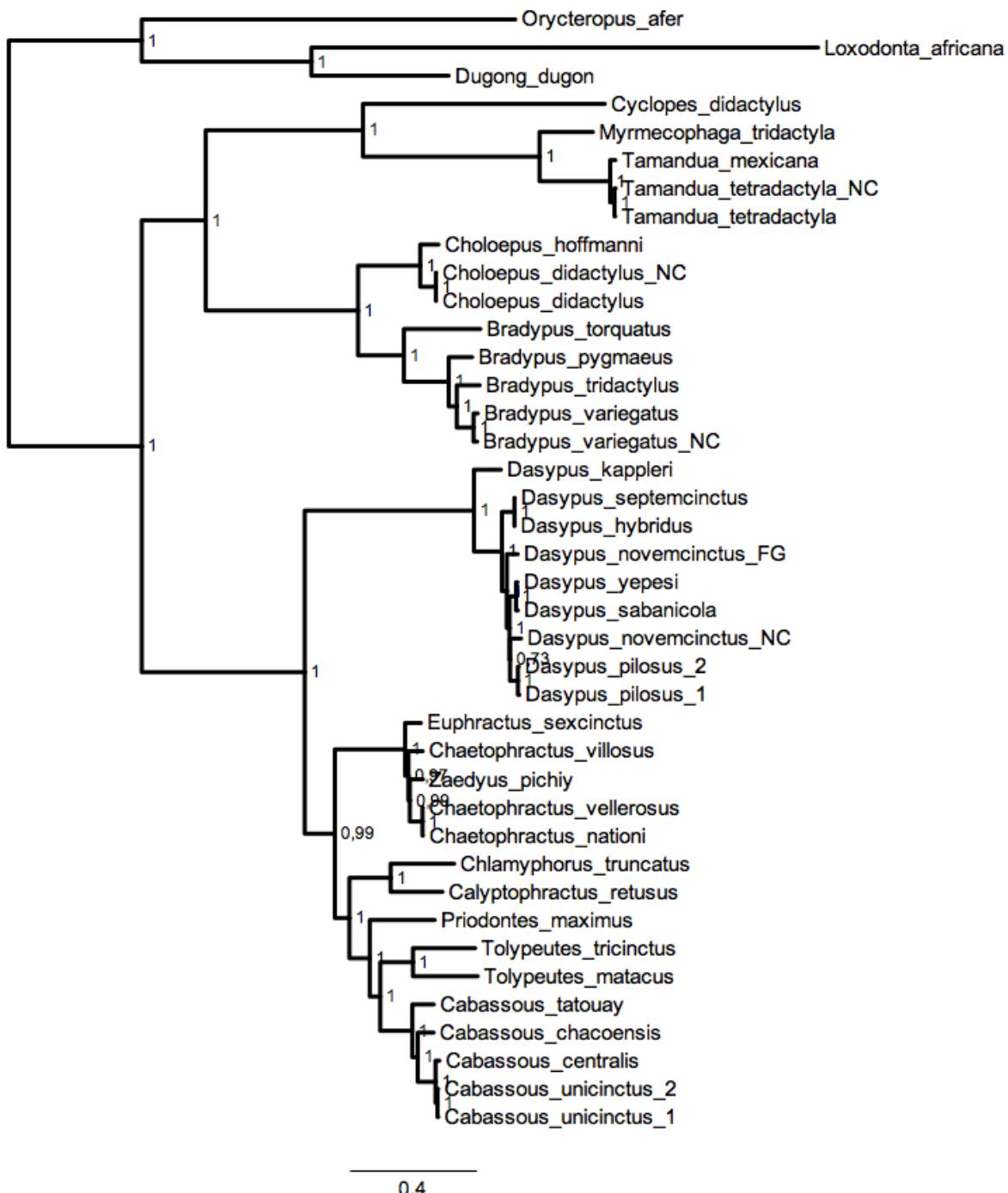
Models	Mode of dependence	NP	logL	AICc	ΔAICc	λ	α	μ	β
λ varies with temperature	Exponential	2	-117.673 ± 0.262	239.760 ± 0.524	0	0.097 ± 0.0023	-0.0912 ± 0.0032	-	-
λ varies with temperature and μ constant	Exponential	3	-117.623 ± 0.258	242.103 ± 0.516	2.344	0.095 ± 0.0023	-0.0699 ± 0.0052	0.015 ± 0.0023	-
λ constant and μ varies with temperature	Exponential	3	-117.917 ± 0.241	242.692 ± 0.482	2.932	0.078 ± 0.0014	-	0.047 ± 0.0032	0.0442 ± 0.0121
Both λ and μ vary with temperature	Exponential	4	-117.581 ± 0.255	244.644 ± 0.510	4.885	0.099 ± 0.0029	-0.0713 ± 0.0062	0.097 ± 0.0356	-0.0355 ± 0.0222
λ varies with temperature	Linear	2	-117.718 ± 0.265	239.850 ± 0.530	0.090	0.086 ± 0.0020	-0.0049 ± 0.0003	-	-
λ varies with temperature and μ constant	Linear	3	-117.627 ± 0.262	242.111 ± 0.524	2.352	0.086 ± 0.0018	-0.0038 ± 0.0003	0.014 ± 0.0021	-
λ constant and μ varies with temperature	Linear	3	-117.904 ± 0.241	242.666 ± 0.482	2.906	0.076 ± 0.0015	-	0.060 ± 0.0035	-0.0048 ± 0.0006
Both λ and μ vary with temperature	Linear	4	-117.554 ± 0.261	244.589 ± 0.523	4.830	0.092 ± 0.0022	-0.0037 ± 0.0003	0.050 ± 0.0056	-0.0038 ± 0.0006

(b) Time-dependence (discrete shift) approach (TreePar; Stadler 2011). Values are means and standard errors calculated over 100 trees taken randomly from the Bayesian dating analyses. NP, number of parameters; logL, log-likelihood; r1, net diversification rate (speciation minus extinction) from the Present to the first shift time (st1); and τ_1 , turnover (extinction over speciation) from the Present to the first shift time (st1). The best-fitting model is a constant birth-death model with no significant rate shift detected.

Models	NP	logL	AICc	ΔAICc	r1	τ_1	st1	r2	τ_2	st2	r3	τ_3	st3	r4	τ_4
Constant bd	2	-97.17 ± 0.241	198.75 ± 0.482	0	0.0275 ± 0.0005	0.6381 ± 0.0136	-	-	-	-	-	-	-	-	-
BD with 1 st	5	-93.59 ± 0.260	199.48 ± 0.520	0.73	-0.2378 ± 0.0464	2.0191 ± 0.2131	39.7 ± 2.52	-2.2963 ± 0.1902	1.8227 ± 0.1952	-	-	-	-	-	-
BD with 2 st	8	-91.42 ± 0.251	205.11 ± 0.502	6.35	-0.2742 ± 0.0625	2.4597 ± 0.3221	19.9 ± 2.16	-0.2079 ± 0.1063	43.5587 ± 41.1184	46.6 ± 2.08	-2.6118 ± 0.1790	2.2113 ± 0.3753	-	-	-
BD with 3 st	11	-89.76 ± 0.251	214.72 ± 0.502	15.97	-0.2634 ± 0.0743	2.1096 ± 0.2710	17.5 ± 1.99	-0.0949 ± 0.1364	44.5790 ± 41.1192	27.8 ± 2.16	-0.2794 ± 0.1383	1.4605 ± 0.3257	48.3 ± 1.96	-2.6356 ± 0.1794	2.4595 ± 0.4616

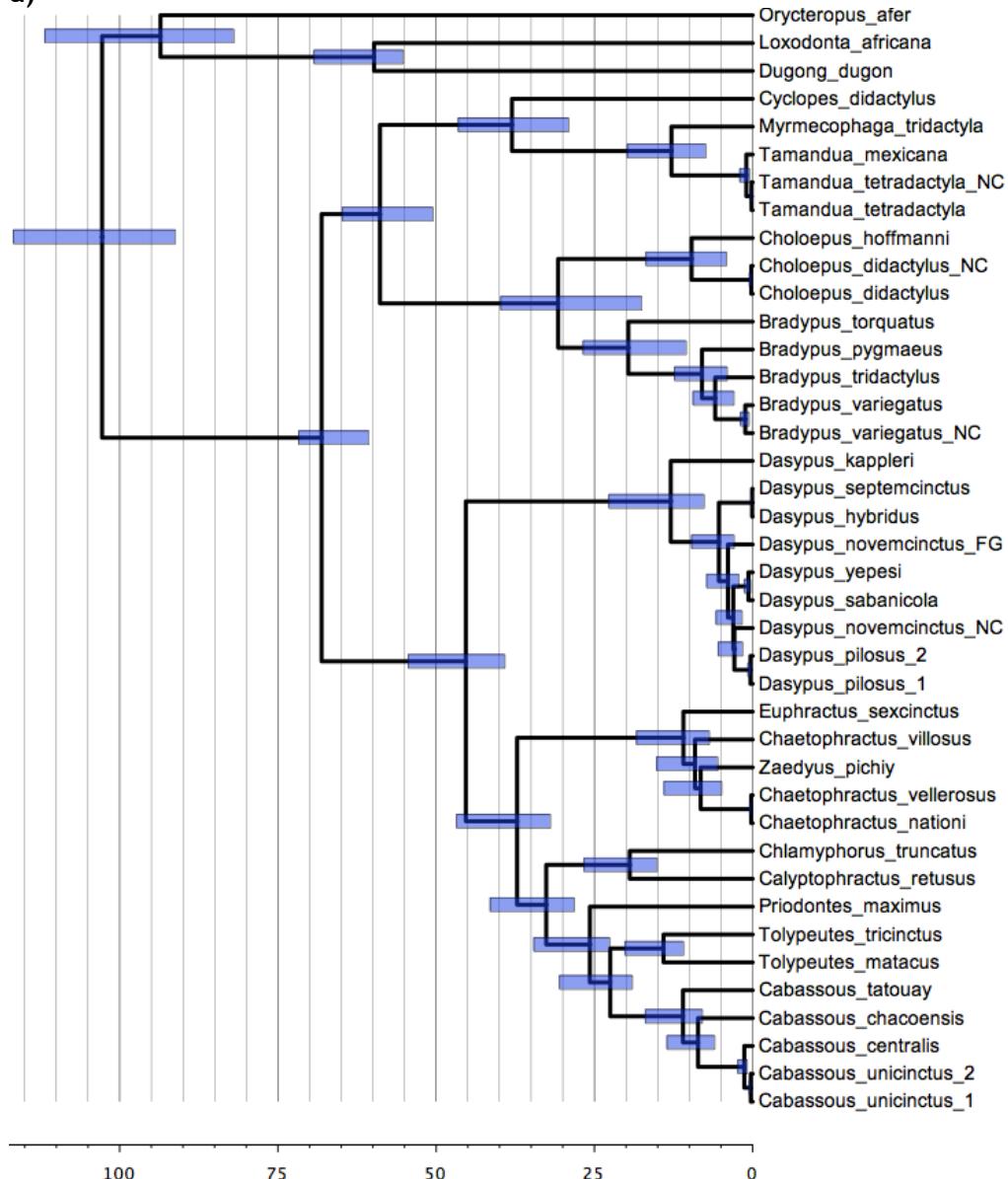
(c) Diversity-dependence approach (DDD; Etienne et al. 2012). NP, number of parameters; logL, log-likelihood; λ , speciation rate; μ , extinction rate; and K, estimated carrying capacity given the maximum likelihood parameters of the model. The best-fitting model is a model in which extinction increases exponentially as the diversity of the clade increases over time.

Current species richness	Models	NP	logL	AICc	ΔAICc	λ	μ	K
32, and all species are sampled	DDL	2	-115.978	236.369	11.347	0.05	-	Inf
	DDL+E	3	-114.200	235.258	10.236	0.2061	0.09245	30.77
	DDX+E	3	-114.689	236.235	11.213	0.32	0.07762	43.92
	DD+EL	3	-113.906	234.668	9.646	0.1009	≈ 0	34.27
	DD+EX	3	-109.083	225.022	0	0.0954	≈ 0	31.39



Supplementary Figure S1: Complete Bayesian consensus tree.

Complete Bayesian consensus phylogram obtained using PhyloBayes under the CAT-GTR-G mixture model. Values at nodes indicate Bayesian posterior probabilities (PP).

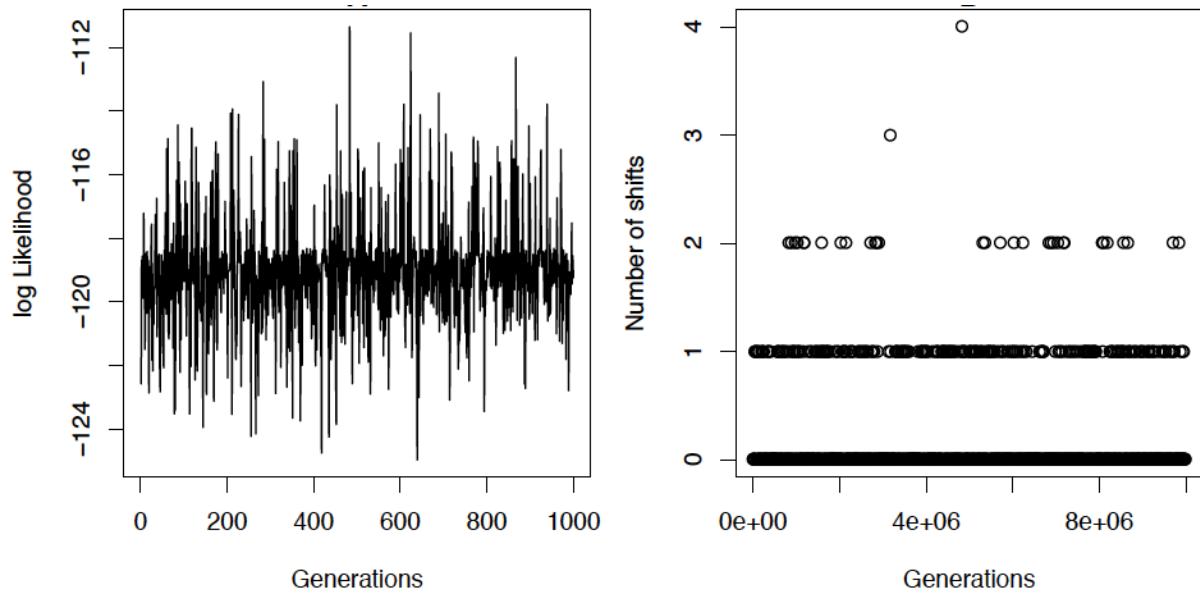
a)**b)**

```
((Orycterus_afer:93.5553, (Loxodonta_africana:59.821, Dugong_dugon:59.821):33.7343):9.2258, (((Cyclopes_didactylus:38.0246, (Myrmecophaga_tridactyla:12.8339, (Tamandua_mexicana:1.0027, (Tamandua_tetradactyla_NC:0.1684, Tamandua_tetradactyla:0.1684):0.8342):11.8312):25.1907):20.8529, ((Choloepus_hoffmanni:9.6461, (Choloepus_didactylus_NC:0.2072, Choloepus_didactylus:0.2072):9.4389):21.0823, (Bradypus_torquatus:19.635, (Bradypus_pygmaeus:7.9961, (Bradypus_tridactylus:5.9217, (Bradypus_variegatus:1.1677, Bradypus_variegatus_NC:1.1677):4.754):2.0743):11.6388):11.0935):28.1491):9.2283, ((Dasypus_kappleri:12.9351, ((Dasypus_septemcinctus:0.0776, Dasypus_hybridus:0.0776):5.259, (Dasypus_novemcinctus_FG:3.9067, ((Dasypus_yepesi:0.6622, Dasypus_sabanicola:0.6622):2.3934, (Dasypus_novemcinctus_NC:2.8906, (Dasypus_pilosus_2:0.3478, Dasypus_pilosus_1:0.3478):2.5427):0.165):0.851):1.43):7.5984):32.3822, ((Euphractus_sexcinctus:10.9399, (Chaetophractus_villosus:9.0975, (Zaedyus_pichiy:8.2149, (Chaetophractus_vellerosus:0.2264, Chaetophractus_nationi:0.2264):7.9884):0.8826):1.8424):26.2721, ((Chlamyphorus_truncatus:19.4188, Calyptophractus_retusus:19.4188):13.1925, (Priodontes_maximus:25.7131, ((Tolypeutes_tricinctus:14.0698, Tolypeutes_matacus:14.0698):8.4247, (Cabassous_tatouay:11.0105, (Cabassous_chacoensis:8.6233, (Cabassous_centralis:1.313, (Cabassous_unicinctus_2:0.2894, Cabassous_unicinctus_1:0.2894):1.0235):7.3102):2.3871):11.4841):3.2184):6.8982):4.6007):8.1053):22.7885):34.6752);
```

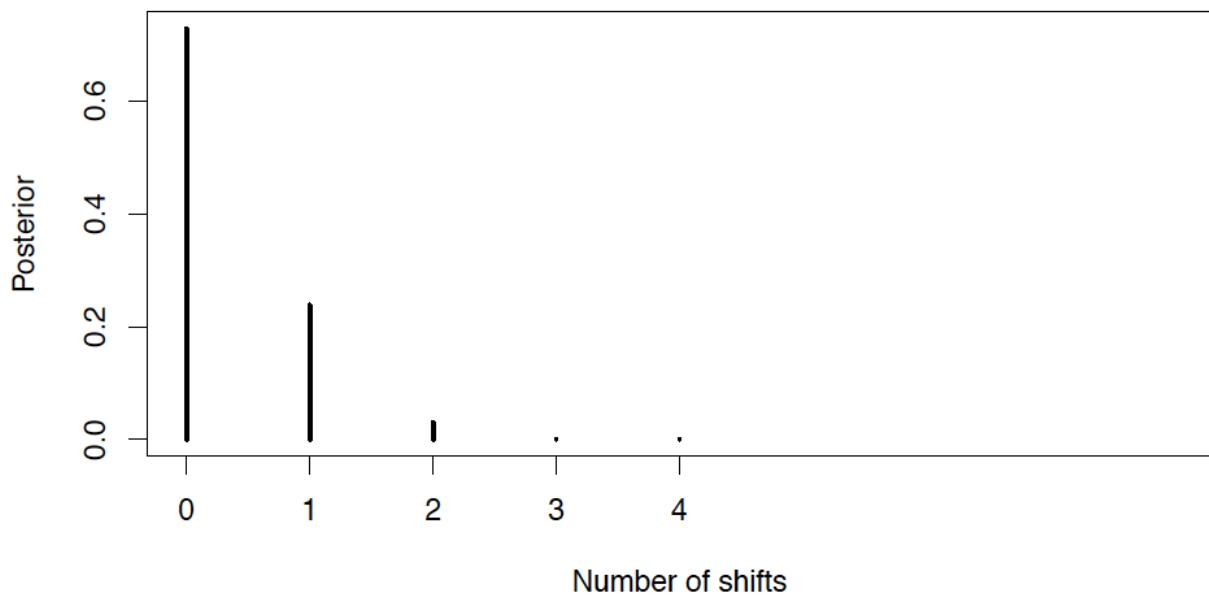
Supplementary Figure S2: Complete Bayesian chronogram.

a) Complete Bayesian chronogram obtained using a rate-autocorrelated log-normal relaxed molecular clock model using PhyloBayes under the CAT-GTR-G mixture model with a birth-death prior on the diversification process, and six soft calibration constraints. Node bars indicate the uncertainty around mean age estimates based on 95% credibility intervals. b) Corresponding chronogram in newick format.

(a)



(b)



Supplementary Figure S3: Detailed BAMM results.

(a) Assessing MCMC convergence of the BAMM run showing that the configuration with no shifts in diversification is the most frequently visited. (b) Posterior probabilities (PP) for the number shifts in diversification rate. The configuration with no shift received a PP of 0.78 whereas the configuration with one shift has a PP of 0.24.